

AD-A107 908

MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB F/G 5/8
THE EFFECTS OF MICROPHONES AND FACEMASKS ON LPC VOCODER PERFORM--ETC(U)
SEP 81 E SINGER F19628-80-C-0002

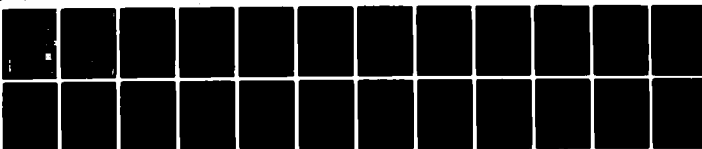
UNCLASSIFIED

TR-584

ESD-TR-81-277

NL

1 of 1
AL
A 07500



END
DATE
FILMED
1 82
DTIC

AD A137908

LEVEL ~~AP~~

12

Technical Report

The Effects of Microphones and Facemasks on LPC Vocoder Performance

Prepared for the Department of the Air Force
under Electronic Systems Division Contract F19620-80-C-0002 by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

E. S. S. S.

DTIC
ELECTE
DEPT

25 September 1981



Approved for public release; distribution unlimited.

81 11 30 058

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, under contract to the Department of the Air Force under Contract F19620-69-C-0002.

This report may be reproduced to satisfy needs of U.S. Government agencies.

The views and conclusions contained in this document are those of the contractor and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Government.

The Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Raymond L. Loiselle

Raymond L. Loiselle, Lt. Col., USAF
Chief, ESD Lincoln Laboratory Project Office

Non-Lincoln Recipients

PLEASE DO NOT RETURN

Permission is given to destroy this document
when it is no longer needed.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

**THE EFFECTS OF MICROPHONES AND FACEMASKS
ON LPC VOCODER PERFORMANCE**

E. SINGER

Group 24

TECHNICAL REPORT 584

25 SEPTEMBER 1981

Approved for public release; distribution unlimited.

LEXINGTON

MASSACHUSETTS

ABSTRACT

↓

The effects of oxygen facemasks and noise cancelling microphones on LPC vocoder performance were analyzed and evaluated. Likely sources of potential vocoder performance degradation included the non-ideal frequency response characteristics of the microphone, the acoustic alterations of the speech waveform due to the addition of the facemask cavity, and the presence of breath noise imposed by the close-talking requirement. It is shown that the presence of the facemask produces a vowel-dependent reduction in the bandwidths of the upper speech formants. In addition, the low frequency emphasis normally associated with small enclosures is shown to occur when a pressure microphone is employed for transduction. Noise cancelling microphones, which are sensitive to the pressure gradient, do not exhibit this effect. Finally, an acoustic tube model of the vocal tract and facemask is presented which predicts the absence of spurious resonances within the frequency band of typical narrowband vocoders. Evidence supporting these assertions is presented based on observed vowel spectra. Evaluations performed using Diagnostic Rhyme Tests indicate that the presence of the oxygen facemask and noise cancelling microphone does not result in a significant increase in the LPC vocoder processing loss.

↑

Dist		
A		

CONTENTS

Abstract	iii
I. Introduction	1
II. Preliminary Considerations	1
III. Analysis of Vocal Tract Termination	4
IV. Facemask Acoustics	8
V. Vowel Spectra	13
VI. DRT Results	16
VII. Conclusions	18
References	20

I. INTRODUCTION

One of the continuing goals of narrowband vocoder research is the development of systems capable of performing successfully in realistic operating environments. Degradations in vocoder performance may arise from many sources: background noise, acoustic distortions, competing speakers, undesirable signal conditioning, channel errors, and so on. The results presented in this report were part of a broader study directed towards the evaluation of LPC-10 vocoder operation in F-15 high performance fighter aircraft. Although it is immediately evident that the ambient noise levels in the F-15 cockpit are quite high, it is also true that considerable noise attenuation is achieved through the use of the oxygen facemask and noise cancelling microphone attached to the pilot's helmet. However, the imposition of a closed cavity such as a facemask over the mouth results in an acoustic modification of the speech signal which could corrupt the waveform and lead to a subsequent loss in vocoder output quality. This report presents the results of a study in which the impact of the oxygen facemask and noise cancelling microphone was assessed through the application of acoustic theory, an extensive examination of vowel spectra, and the use of Diagnostic Rhyme Testing.

II. PRELIMINARY CONSIDERATIONS

The project was begun with some general notions as to the nature of the acoustic distortions introduced through the coupling of the oxygen facemask and noise cancelling microphone to the vocal tract. The acoustic effects were presumed to result in modifications to the signal which could be broadly classified as speech-dependent and speech-independent. The low frequency emphasis of the speech signal that results from covering

the mouth with a closed cavity is one example of a speech-independent phenomenon induced by the facemask. Speech-dependent effects were thought to involve fundamental alterations in the formant structure of the natural speech signal and might include the introduction of additional resonances. Since spurious resonances appearing in the frequency band of the vocoder could easily lead to the failure of a 10th order linear prediction model, considerable attention was given to the potential need for a higher order LPC model in a facemask environment.

The choice of microphone and the design of its frequency response characteristic appeared to be another possible source of vocoder performance degradation. Because of the presence of the bass boost phenomenon associated with facemasks, pressure microphones designed for use inside facemasks generally have a tapered low end response to counteract the boost introduced by the mask. The pressure gradient (noise cancelling) microphones employed by the Air Force for use in operational environments invariably exhibit a low frequency rolloff of 6 dB/octave below about 1000 Hz. An example of such a characteristic is shown in Fig. 1 which illustrates the frequency response of a typical M101 gradient microphone designed for use in the MBU-5/P oxygen facemask. Experience has shown that both the LPC modelling process and pitch detection algorithms benefit from signal conditioning designed to keep the average speech spectrum as flat as possible. Deviations from the ideal spectrum can lead to deficiencies in the parameter extraction process which will ultimately be manifested as a reduction in speech intelligibility at the synthesizer.

Another aspect to the use of microphones inside facemasks is the

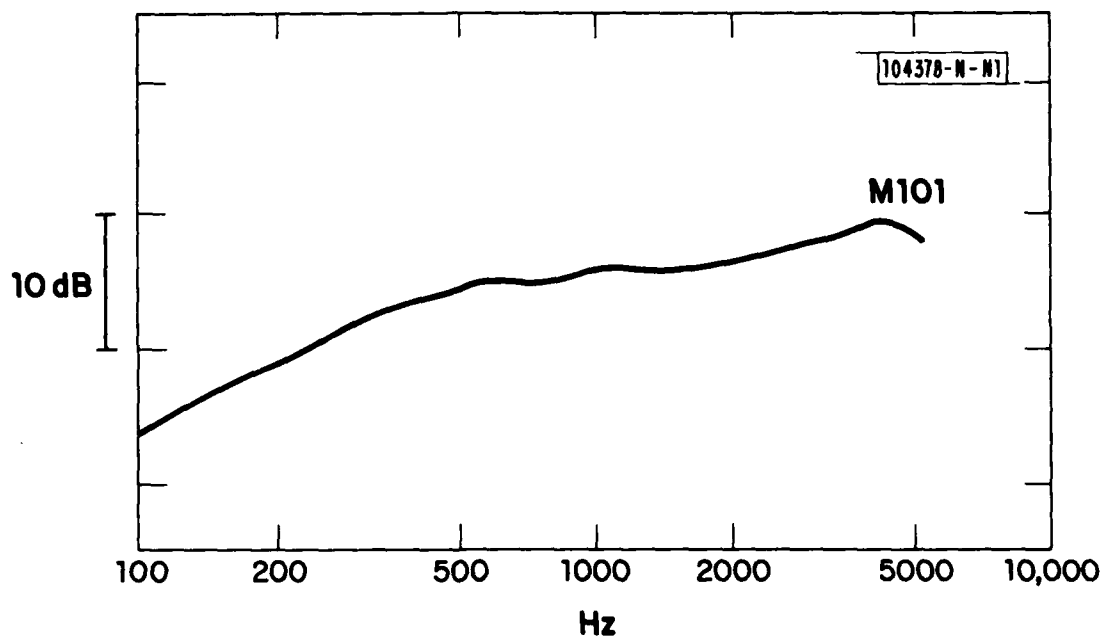


Fig. 1. Frequency response of the M101 pressure gradient microphone.

close talking requirement. The proximity of the microphone to the talker's lips introduces abnormally large bursts of energy into the speech signal, particularly during the articulation of plosives and fricatives. The addition of breath noise to the speech is especially pronounced in microphones with a flat low frequency response. Although the intelligibility of the unprocessed speech is minimally affected by breath noise, the effects on narrowband vocoders can be particularly damaging. Studies have shown that the use of a foam windscreen in conjunction with boom-mounted, close talking microphones can yield a considerable improvement in vocoder speech intelligibility and is a virtual necessity in microphones with a good low end response [1].

III. ANALYSIS OF VOCAL TRACT TERMINATION

Many of the prominent features observed in speech spectra can be derived from a model in which the vocal tract is represented by a single uniform lossless acoustic tube closed at one end (the glottis) and open at the other (the lips). For a tube 17 cm long the natural frequencies occur at approximately 500 Hz, 1500 Hz, 2500 Hz, etc. If the vocal tract were entirely lossless, speech formants would be of infinite amplitude. Losses introduced at the glottis, cavity walls, and termination act primarily to increase the bandwidths of the formants.

The speech-independent effects produced by the facemask and the interaction between the mask and the microphone are best understood using a simple lumped-parameter representation of the voice production mechanism. The model shown in Fig. 2a depicts the vocal tract as a constant volume velocity source with a termination representing the effects of radiation at the lips. The inductance represents the opening of the lips into the air and accounts for the spherical wavefronts of the

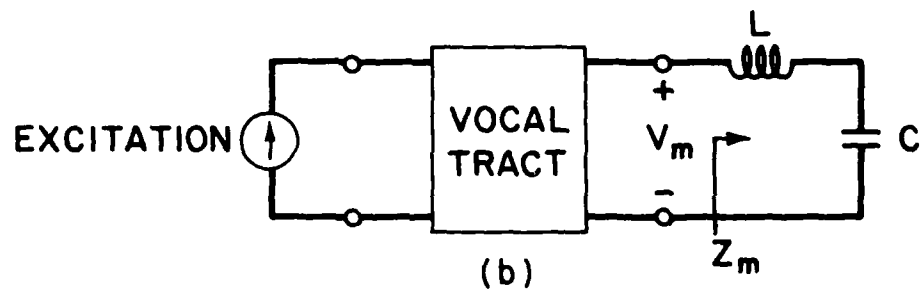
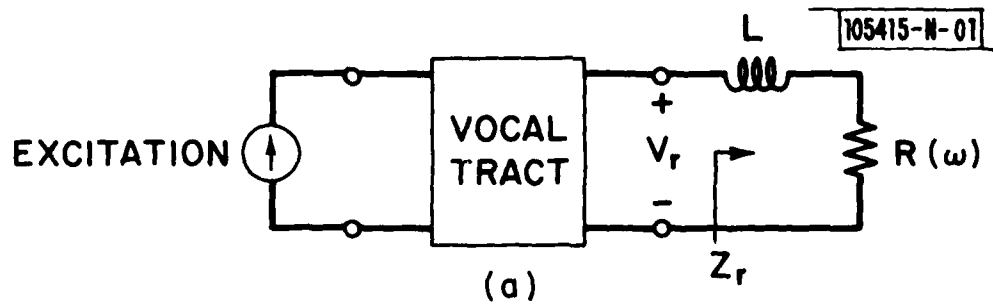


Fig. 2(a-b). Speech production models with (a) open air termination and (b) facemask termination.

sound produced at the lips while the frequency dependent resistance accounts for the radiation of energy into the air. Because the termination is primarily resistive at higher frequencies it can be shown that its major effect is a broadening of the upper speech formants [2].

A low frequency model for the facemask condition can be obtained by replacing the termination shown in Fig. 2a with that of Fig. 2b. The inductance once again models the effects of the lip opening while the capacitance represents the volume of the facemask. Because the facemask is nearly closed, considerably less energy is radiated under this condition and the effects of energy loss are therefore ignored. This model predicts a narrowing of the higher frequency formants relative to those observed in the unconstrained situation, particularly for those sounds accompanied by large energy levels. Of course, the vocal tract is not an ideal source and the change in termination will cause the natural frequencies to shift from their normal values and could lead to the appearance of additional formants. The nature of these modifications to the formant pattern is best understood using acoustic tube theory and will be discussed in the next section.

The effect of the facemask on the output of a pressure-sensitive microphone can also be examined with the aid of Fig. 2. Assuming the vocal tract to be a high impedance source, the ratio of the pressure response at the lips for the facemask condition relative to the response in the open is given by

$$\frac{V_m}{V_r} = \frac{Z_m}{Z_r} ,$$

where Z_m and Z_r represent the termination impedances. From the sketch of this function shown in Fig. 3 one may observe the 12 dB/octave rise

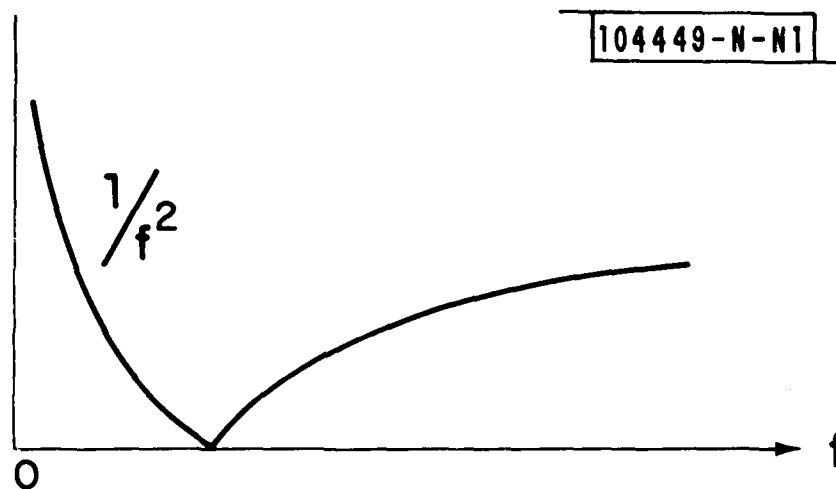


Fig. 3. Pressure response in facemask relative to pressure response in open air.

in the pressure response introduced by the presence of the facemask. Pressure microphones used inside facemasks employed by the British have frequency responses specifically tailored to compensate for this effect.

The pressure gradient microphone employed inside the facemask for its noise cancellation capability is designed to measure the difference in pressure at its front and back ports. The operation of the pressure gradient microphone can be understood by distributing the effects of the lip inductance and determining the voltage difference across one section, as illustrated in Fig. 4. It is apparent that the voltage difference is dependent on the current flow or volume velocity rather than on the voltage. If the vocal tract is once again treated as a high impedance source, the output of the pressure gradient microphone is approximately independent of the particular termination applied to the tract. Thus, while the effects of vocal tract/facemask coupling will be evident in the response of the gradient microphone, no overall spectral shaping will be introduced by the mask and no compensation using the microphone frequency response is necessary.

IV. FACEMASK ACOUSTICS

An understanding of the interaction between the facemask and vocal tract can be obtained by representing the structures as a concatenation of two acoustic tubes each open at one end and closed at the other, as shown in Fig. 5. The acoustic impedances looking into the open ends of the individual tubes are given by

$$Z_1 = -j \frac{\rho c}{A_1} \text{ctn } (\omega l_1/c)$$

and

$$Z_2 = -j \frac{\rho c}{A_2} \text{ctn } (\omega l_2/c)$$

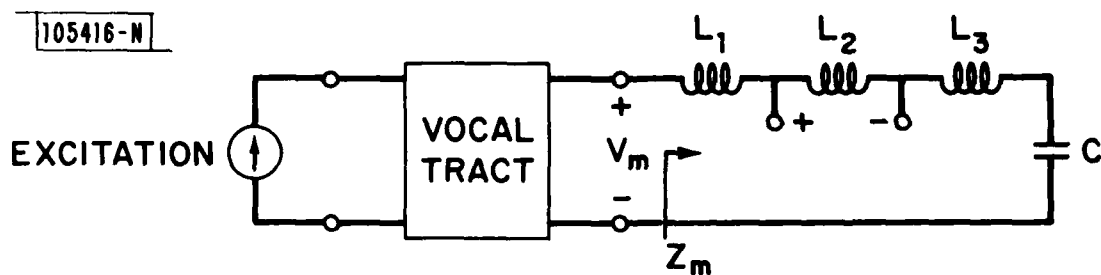


Fig. 4. Model for pressure gradient microphone response mechanism.

where A_1 and A_2 are the cross-sectional areas of the tubes, ℓ_1 , and ℓ_2 are the tube lengths, c is the velocity of sound, and ρ is the density of air [2]. The natural frequencies of the system occur when

$$Z_1 + Z_2 = 0$$

or

$$\text{ctn } (\omega \ell_1 / c) = -\frac{A_1}{A_2} \text{ctn } (\omega \ell_2 / c). \quad (\text{Eq. 1})$$

The left and right hand sides of Eq. 1 are shown, respectively, by the solid and dashed lines of Fig. 6. Open circles represent the natural frequencies of the unconstricted vocal tract while the solid circles indicate the position of the natural frequencies of the modified acoustic system. The curves were drawn with the tract and mask dimensions given in Fig. 5 which crudely represent the acoustic cavities. The figure indicates that the first additional resonance occurs in the vicinity of $2c/4\ell_2$ or about 5800 Hz, well beyond the upper limit of most narrowband vocoders. Lengthening the facemask lowers the frequency of the first extra resonance, while increasing the cross-sectional area of the facemask increases the shift of the natural frequencies relative to their normal positions. Consequently, the change in the speech formant frequencies depends on both the length and the cross-sectional area of the facemask. The curves in Fig. 6 suggest that the mask would tend to increase the frequency of the first formant and to affect the next three formants to a much lesser extent. Of course, the time-varying cross-sectional area of the vocal tract differs substantially from that of the ideal acoustic tube and thus observations made on real speech will yield results that change accordingly. It is of interest to note that Morrow [4] and Morrow and Brouns [3] observed an upward shift in the first speech formant in their investigations.

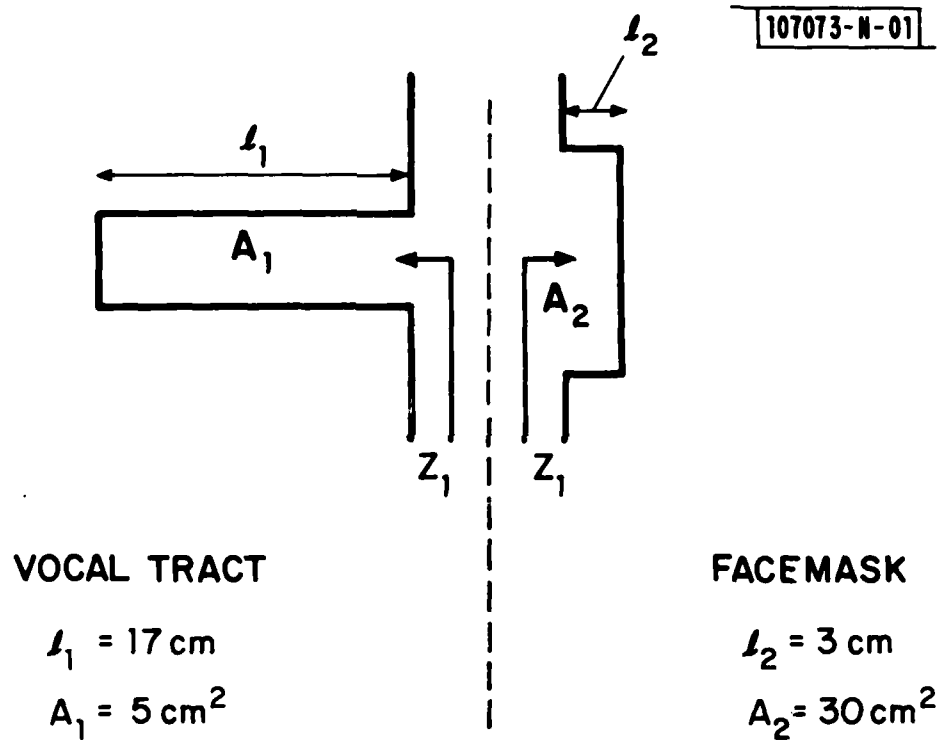
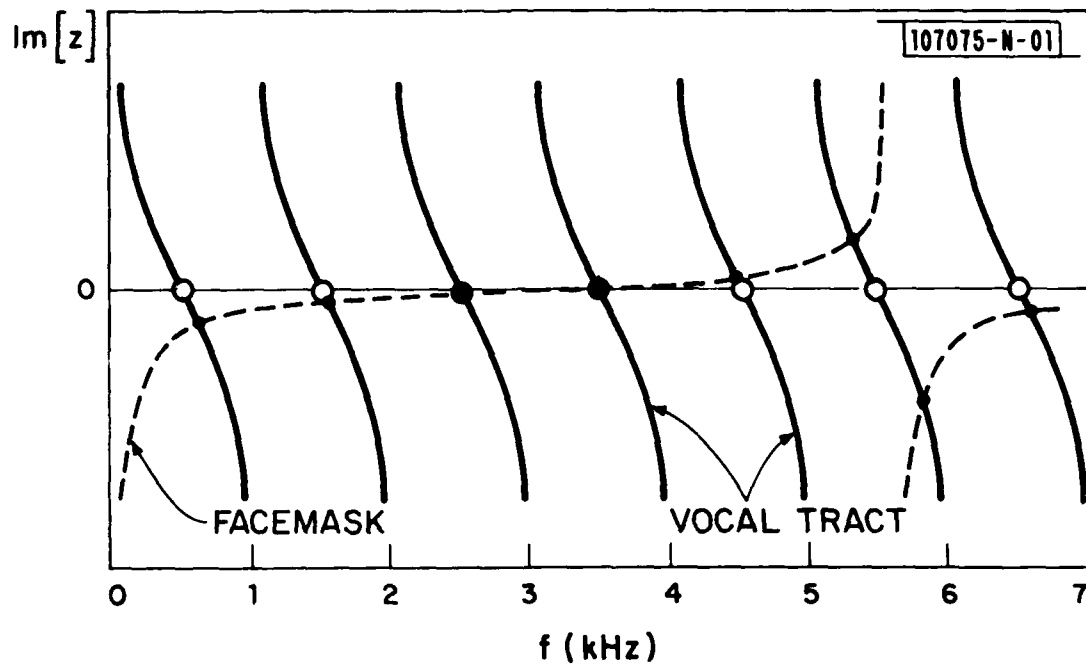


Fig. 5. Acoustic tube representation of the vocal tract and facemask.



- UNCONSTRICTED VOCAL TRACT
- VOCAL TRACT WITH FACEMASK

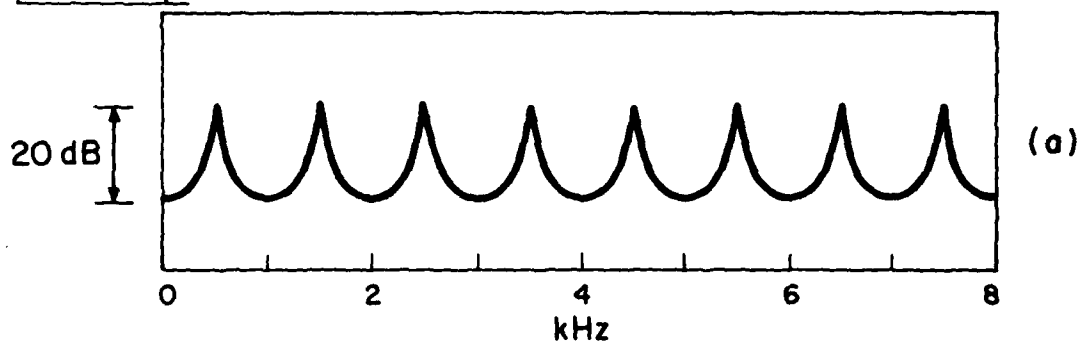
Fig. 6. Effect of the facemask on the vocal tract natural frequencies.

Evidence supporting this theory was obtained using the facilities of the Speech Communication Laboratory at MIT. A computer simulation was used to calculate the magnitudes of the transmission functions of an unconstricted uniform acoustic tube and of the same tube terminated with a closed section representing the facemask. The transmission function shown in Fig. 7a is that of the unconstricted lossy acoustic tube and clearly depicts formants occurring at $(2n+1)*500$ Hz. The transmission function of a lossy two-tube system is shown in Fig. 7b and illustrates the predicted upward shift of the first formant. It can be seen that the first resonance obviously associated with the terminating tube does not appear within the frequency range of the vocoder.

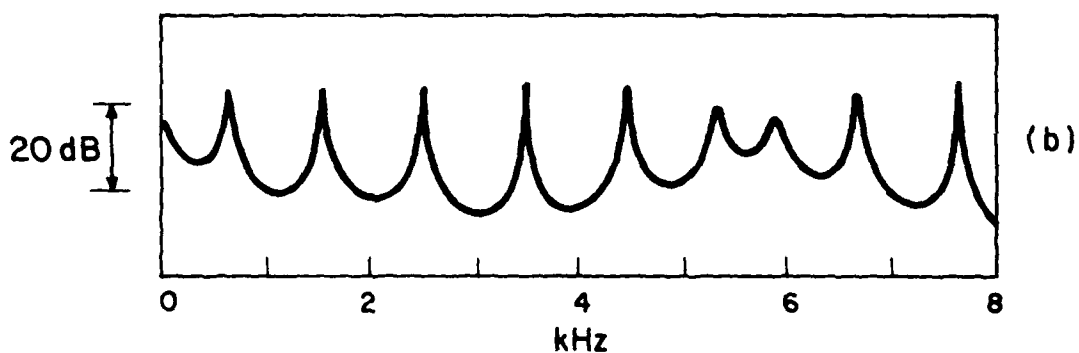
V. VOWEL SPECTRA

The validity of the theoretical results presented above was evaluated through an extensive examination of vowel spectra obtained from speech produced using the oxygen facemask and noise cancelling microphone. The spectra of the vowel portions of the word "fad" as recorded using both a high quality dynamic microphone and the facemask and noise cancelling microphone are shown in Fig. 8. It is apparent that no resonances are present in the speech produced with the facemask that are not observed in the open condition. No such artifacts have in fact appeared in any of the vowel segments which have been examined. It is also evident that although the formant frequencies have not been significantly altered, the bandwidths of the higher formants have been reduced considerably. This effect was most noticeable in the vowels in "fed," "fad," and "fod" but was not observed in others. It can also be seen that the frequency response of the M101 gradient microphone as measured in open conditions (Fig. 1) is reflected in the spectrum obtained using the facemask.

107076-N-02



UNIFORM ACOUSTIC TUBE



UNIFORM ACOUSTIC TUBE
TERMINATED WITH FACEMASK

Fig. 7(a-b). Transmission magnitude spectra for (a) uniform acoustic tube and (b) uniform acoustic tube terminated with facemask.

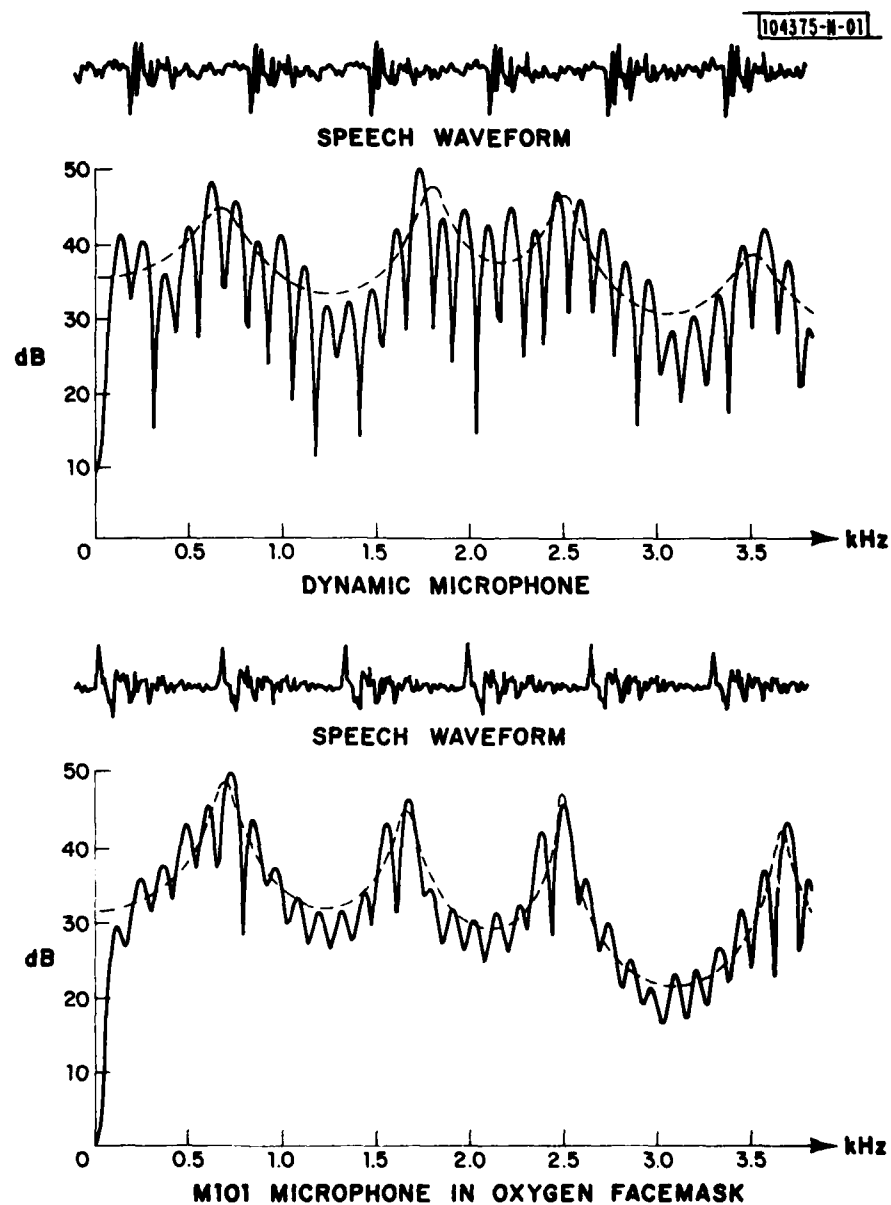


Fig. 8. Effects of the microphone and facemask on the speech spectrum.

Of greater significance is the ability of LPC-10 to model the spectra of the speech produced by the facemask. The LPC spectral fits for the two conditions are shown by the dashed lines in Fig. 8. It appears that the ability of LPC-10 to model speech produced through the facemask and gradient microphones is comparable to its ability to model speech obtained using a high quality microphone in the open. Consistent with the theory presented earlier, no mechanism inherent to the facemask has been identified which would obviously produce a breakdown in the 10th order spectral modelling process.

VI. DRT RESULTS

A series of DRTs was conducted to obtain an objective assessment of the performance of the Lincoln LPC-10 [5] vocoder in the F-15 environment. Three speakers, all former or active Air Force fighter pilots, were used as subjects. Each speaker was required to read the DRT word lists while wearing a helmet containing a facemask and noise cancelling microphone.

The results of the DRT are shown in Fig. 9 with the bars indicating the maximum, minimum, and 3-speaker average scores. The first section of the graph illustrates the performance of the LPC-10 algorithm under normal conditions using a high quality dynamic microphone. The processing loss through LPC-10 for this reference condition is about ten DRT points. The second section of the graph shows the effects of the facemask and microphone on both the processed and unprocessed speech signals. It appears that the introduction of the facemask has produced a significant intelligibility loss (3.5 points) in the unprocessed speech. However, the average score for the processed speech shows that the processing loss incurred through LPC-10 is only slightly greater than that observed under ideal conditions.

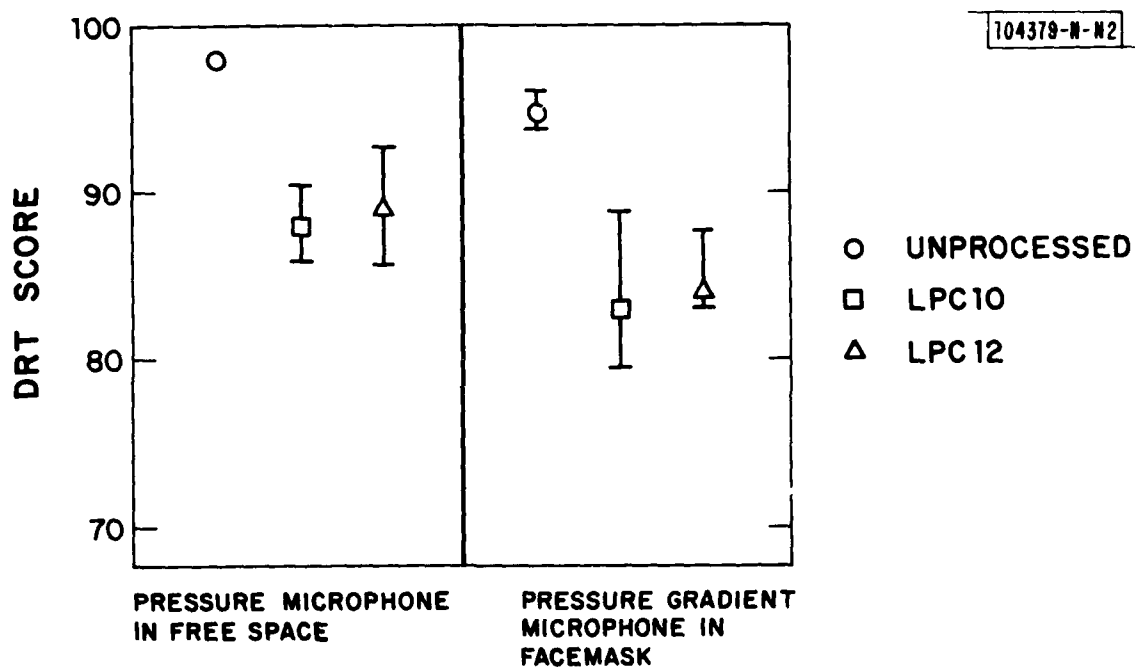


Fig. 9. 3-speaker DRT evaluation of effects of gradient microphone and facemask.

In addition to obtaining DRT results for the standard 2.4 kbps LPC-10 system, testing was also conducted using a fixed-point 12th order LPC algorithm. For this system coding was used only for the pitch and energy parameters; the 12 uncoded reflection coefficients were transmitted to the synthesizer directly. The scores for LPC-12 shown in Fig. 9 illustrate that no significant gain in performance is achieved by imposing a higher order spectral model on the signal and that the presence of the facemask alone does not appear to introduce additional complexity to the waveform. Of course these conclusions must be tempered by the fact that the data is based on a limited speaker population which exhibited a fairly wide range of processed speech scores.

Finally, a more limited set of tests was conducted to determine the effects of windscreens on LPC-10 performance when used in conjunction with the M101 microphone inside the facemask. Three conditions were examined, two noise conditions each involving a single speaker and a noise-free condition using two speakers. In none of these situations did the performance of the LPC-10 algorithm benefit from the use of a windscreen. Informal listening corroborated these findings and indicated that the breath noise was of no particular concern using the present facemask/microphone combination.

VII. CONCLUSIONS

Although the results presented in this report on the impact of facemasks and noise cancelling microphones on LPC vocoder performance are based on a fairly limited talker base, it is nevertheless possible to draw some conclusions from the data. The DRT scores demonstrate that the unprocessed speech signal produced using the facemask undergoes a considerable reduction in intelligibility relative to normal speech.

Several mechanisms have been identified which may contribute to this loss; specifically, a narrowing of the upper formant bandwidths and an upward shift in the lower formant frequencies. However, both theoretical and experimental evidence indicate that the complexity of the speech waveform is not increased by the imposition of the facemask and thus the validity of an LPC-10 spectral fit is not compromised. It was also shown that the low frequency emphasis associated with the application of a closed cavity to the vocal tract is not present when a pressure gradient microphone is employed. Consequently, the frequency response characteristic of a gradient microphone in an open condition will be observed in a facemask environment as well. Since the presence of a low frequency rolloff tends to reduce vocoder performance, it would be beneficial to restore the low frequency response in operational noise cancelling microphones. This improved low end response is likely to magnify the problems of breath noise, however, and the addition of foam windscreens may be necessary.

REFERENCES

- [1] "ANDVT Microphone and Audio System Optimization Study-Final Report," Ketron No. 1485, Ketron, Inc., Wayne, PA (February 1980).
- [2] J. L. Flanagan, Speech Analysis, Synthesis and Perception, 2nd. ed., (Springer-Verlag, New York, 1972).
- [3] C. T. Morrow and A. J. Brouns, "Speech Communication in Diving Masks. Part II: Psychoacoustic and Supplementary Tests," J. Acoust. Soc. Am. 50, 1 (1971).
- [4] C. T. Morrow, "Reaction of Small Enclosures on the Human Voice. Part II: Analyses of Vowels," J. Acoust, So. Am. 20, 4 (1948).
- [5] E. M. Hofstetter, P. E. Blankenship, M. L. Malpass and S. Seneff, "Vocoder Implementations of the Lincoln Digital Voice Terminal," Proc. EASON '75, Washington DC, September 29-October 1975.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ESD-TR-81-277	2. GOVT ACCESSION NO. AD A3007-2	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Effects of Microphones and Facemasks on LPC Vocoder Performance		5. TYPE OF REPORT & PERIOD COVERED Technical Report
7. AUTHOR(s) Elliot Singer		6. PERFORMING ORG. REPORT NUMBER Technical Report 584
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lincoln Laboratory, M.I.T. P.O. Box 73 Lexington, MA 02173		8. CONTRACT OR GRANT NUMBER(s) F19628-80-C-0002
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Systems Command, USAF Andrews AFB Washington, DC 20331		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element Nos. 27417F, 64754F and 28010F Project Nos. 2283, 411L and 2264
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Electronic Systems Division Hanscom AFB Bedford, MA 01731		12. REPORT DATE 25 September 1981
		13. NUMBER OF PAGES 28
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES None		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) noise cancelling microphones acoustic analysis of speech facemasks degraded speech linear predictive coding		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The effects of oxygen facemasks and noise cancelling microphones on LPC vocoder performance were analyzed and evaluated. Likely sources of potential vocoder performance degradation included the non-ideal frequency response characteristics of the microphone, the acoustic alterations of the speech waveform due to the addition of the facemask cavity, and the presence of breath noise imposed by the close-talking requirement. An acoustic tube model is used to demonstrate that the facemask does not introduce spurious resonances within the frequency band of typical narrowband vocoders. Further analysis indicates that the low frequency emphasis normally associated with small enclosures is not present when pressure gradient microphones are employed. Evidence supporting these assertions is presented based on observed vowel spectra and the results of Diagnostic Rhyme Tests.		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

**DATA
FILM**